

Flores, M. A. & Derrington, M.L. (2017) School principals' views of teacher evaluation policy: lessons learned from two empirical studies, *International Journal of Leadership in Education*, vol. 20, nº 4, pp. 416-431

School principals' views of teacher evaluation policy:

Lessons learned from two empirical studies

Abstract

Teacher evaluation accountability policies are emerging world-wide. This paper examines principals' perceptions of teacher evaluation in terms of commonalities and differences arising from two research projects conducted in Portugal and in the US. Perceptions of school principals in regard to a new policy on teacher evaluation as well as its perceived effects at school and conditions for its implementation will be analysed. Findings point to the challenges and successes of coping with mandated accountability measures in two different contexts. Principals balanced perceived conflicting goals, sought maintenance of positive teacher relationships and school culture, and managed the tensions of policy implementation and the making sense of its effects at school. Implications of the findings are discussed.

Keywords: principal perceptions, teacher evaluation, policy design and implementation

Introduction

Accountability, efficiency, and competition are hallmarks of educational leadership's current policy climate (Magno, 2013). While national policy makers debate mandates' merits, principals hold the key to successfully implementing this complex and dynamic process at the local level

(Retallick & Fink, 2002). As principals confront new accountability policies, perhaps one of the most challenging recently has been teacher appraisal or evaluation systems. Significantly changing an evaluation system through legislation related to career progression, compensation and tenure affects teachers' work and lives. As principals attend to teachers' professional development needs and to organisational requirements for accountability, the supervisory relationship between principal and teachers can also be affected (Tuytens & Devos, 2010). To successfully traverse the teacher accountability policy terrain, principals must find ways to merge both teacher and system needs.

Research has demonstrated the influence of school leadership in fostering teacher learning and development (AUTHOR 1, 2004) as well as the key role and the complexity of school leadership in implementing policy initiatives by making sense of them (AUTHOR 1, 2009, 2010; Vekeman, Devos & Tuytens, in press). The purpose of this paper is to analyse key themes arising from research on teacher evaluation policy implementation taking into account the principals' views. In particular, this paper draws upon studies in which both authors have been involved over the last few years. The goal is to explore the lessons learned in developing accountability policies of teacher evaluation and managing their implementation in schools taking into account the aims of the policies, school principals' sense-making of them and teachers and schools' needs and expectations.

World vision and local view

As Mango (2013, p. 3) argues, 'We are currently in the midst of massive educational reform, globally'. Global influences affect countries at the same time local change processes are underway (Mango, 2013). Thus, globalisation has escalated demands for accountability policies

while exposing local cultural differences and implementation of those policies in practice (Berzano & Grimaldi, 2013; Magno, 2013). Mango (2013) stresses that globalisation and localisation of accountability policies is not mutually exclusive. Thus, a broad search for solutions extending beyond one's local or national experience is required. Educators develop a keener perspective and find shared solutions by studying a problem with others in different contexts. Understanding educational reforms through the lens of other countries' policies broadens and extends our thinking and expands the number of potential solutions.

International educational reforms have taken many forms and one of the most widespread is the emphasis on quality assurance of the teaching force. Policy makers in many countries espouse the belief that improving teaching quality will lead to educational success (Barzano & Grimaldi, 2013). More specifically, reformers have promoted teacher evaluation as the best vehicle for judging quality and assuring that every classroom has a highly qualified teacher. Darling-Hammond (2013) claimed a breakdown of teacher supervision and evaluation processes had occurred. Supposedly the evaluation process was designed to improve instruction but improvement was not evident. Darling-Hammond and Marshall (2009) identified the lack of time supervisors spend in classrooms followed by ineffective or substantive feedback to teachers to use for the improvement instruction. Signaling a need for change in the evaluation process, Marshall (2009) asserted that evaluation had become a routine and noted that a ritualistic approach to the evaluation process is problematic especially if impact on student learning is not considered by the evaluator. Reformers claimed that evaluation procedures had become an ineffective formality disconnected from student outcomes and consequently a process to more vigorously judge teacher performance was warranted (Darling-Hammond, 2013, Marshall, 2009, Weisberg, Sexton, Mulhern, & Keeling, 2009). Consequently, principals and other school leaders bore the responsibility for implementing a new and more rigorous evaluation system.

Various competitive and performance-related teacher evaluation policies have been introduced and pressure to implement them has been evidenced from international organisations (OECD,

2009). Acknowledging that similar accountability policies might generate different local results, Mango (2013) states that leaders across contexts “face similar challenges, demands and are increasingly interconnected through international agencies and testing regimes” (p. 203).

It is through the lens of a global problem/local solution quest that we examine the two studies in this paper for what we can learn from the differences and commonalities. A similar comparison was undertaken within a study of a Northeast US state and Beijing, China (Sun, Youngs, Yan, Chu & Zhao, 2012) and concluded that principal evaluation can be a powerful policy instrument to promote instructional leadership. Particularly when examining teacher evaluation systems, learning from other countries is appropriate (Liu & Zhao, 2013). Thus the analysis of our studies is useful given our shared history, challenging present and hopeful future.

It is imperative that the role of principals is also examined when teacher evaluation is studied. Principals are “situated precisely at the accountability nexus between education policy and practice” (Mango, 2013, p. 179). In this regard, Vekeman, Devos and Tuytens (in press, p. 2), argue that “the principal’s agenda of policy implementation may be either consistent or conflictual with the expectations of teachers about the implementation of teacher evaluation policy in their school”. This paper adds to existing research literature on the importance of principal leadership in implementing a new teacher evaluation system. It contributes to a better understanding of how principals’ views of a new policy for teacher evaluation shape its implementation. Although the contexts are different between our two countries in some details, the implementation of a new teacher evaluation system resulted in similar principal concerns and perceptions. Thus by contributing to the collective understanding of evaluation reforms and the leadership challenges of policy implementation, we might also identify potential pitfalls to avoid.

Key features on teacher evaluation in the two contexts

In this section, examples from the two teacher evaluation policies are presented as well as the principal status in both countries (see Table 1). Race to the Top (RTTT) in the United States is the Obama administration’s federally funded competitive grant program launched in 2010. It has prompted rapid changes in many states’ teacher evaluation systems. The Southeastern US state in which this study was conducted was awarded nearly \$500 million, due in part to new legislation and new State Board of Education’s policies complying with the grant’s stringent teacher evaluation guidelines. The changes included reducing teacher observations from 100% of each summative evaluation to 50%. Value-added test scores now account for 35% of the outcome; the remaining 15% is drawn from a large set of achievement options. After the end of every school year, each teacher receives a summative, one-to-five “composite” score. The accountability legislation links these teacher scores to school districts’ retention or separation decisions.

The previous evaluation policy for professionally licensed, tenured teachers required one complete evaluation cycle with one observation every five years. With Race to the Top funding, however, the new policy requires four complete evaluation cycles and observations, covering a minimum of 60–90 minutes yearly, depending on the teacher’s experience. Feedback after observations must now follow specified time limits.

Prior to 2011-12, a set of planning, instructional, and professional indicators were used in evaluations. However, detailed rubrics were not integral to either the observations or the evaluation procedures. In contrast, under the RTTT grant, specific and lengthy scoring rubrics are required with teacher performance rated on a one-to-five point scale.

INSERT TABLE 1 NEAR HERE

In Portugal, a feature of teacher performance appraisal is the large number of legal texts produced to regulate the process since the January 2007 Teacher Career Statute's publication. With the new government elected in June 2011, a new legal framework (Decree n° 26/2012) for teacher appraisal was published in February 2012. Underpinning the new model was the need to simplify the appraisal process yet retain a rigorous system. Stipulations included the following:

- a) longer periods for the appraisal process replacing the previous two-year timeline; b) recognition of the appraisal's formative dimension; c) external teacher appraisals under specific conditions. A qualified teacher teaching the same subject in a different school conducts an external evaluation focusing on the same criteria. Classroom observation and external appraisal are mandatory except for the following: a) teachers aspiring to obtain the top rating of excellent; b) teachers in their second or fourth career stage; c) teachers previously rated as insufficient; and d) teachers in their probationary year. The internal and external appraisal components comprise 60% of the summative evaluation. Additional appraisal components include in-school activities and connections with the community (20%) as well as in-service education and professional development (20%).

In each school or cluster of schools, a committee for coordinating teacher-performance appraisals conducted the process along with an appraiser designated by a faculty member in the same department as the appraisee. The evaluation instruments included a self-evaluation report and the global form.

Brief overview of the research projects

A detailed description of the empirical studies in which both authors have been involved over the last few years is beyond the scope of this paper. In this section, a brief overview is presented in order for the reader to learn about the nature, scope and aims of both research projects whose details can be found elsewhere (AUTHOR 1, 2010; 2012; 2014; AUTHOR 2, 2014).

The study in the Southeastern United States used interview questions in order to examine principals' views of the new system for teacher performance appraisal. Each principal was interviewed for approximately 50 minutes. Interviews were audio recorded with permission and transcribed verbatim. Analysis began with examining the interview data source (Creswell, 2003), to obtain a general sense of the information. A coding process organised the questions and responses into categories, using the lens of principal leadership and teacher evaluation. Topics were grouped or "chunked" (Creswell, 2003; Johnson & Christensen, 2004) according to the interview questions. Themes were identified during both the coding process and a subsequent data review. QDA Miner was also used in the analysis for generating data to derive frequency of mentions. To ensure trustworthiness, peer debriefing was used throughout the process (Creswell, 2003). In total, 14 principals participated in this study.

In Portugal, a voluntary questionnaire was e-mailed to school principals throughout the country. Of the 400 responses, 134 (33%) were received between October and November of 2011. Drawing upon earlier work (AUTHOR 1, 2009, 2012) the questionnaire included both closed and open-ended questions covering the following main dimensions: a) motivation and job satisfaction; b) purposes and focus of teacher appraisal; c) key features of the appraisal system; d) the implementation process of teacher appraisal (procedures, appraisers, criteria and instruments); and e) perceived effects of the new policy in schools. Quantitative data were

analysed using SPSS (19.0). Open-ended questions were inductively analysed and substantive themes were identified. This is part of a broader research project which also includes teacher data with the goal to analyse different aspects of teacher appraisal policy as well as its perceived effects on schools and teachers' professional development (see AUTHOR 1, 2012).

Bearing the research context in mind as well as existing international research literature the aim of this paper is to explore key themes and to discuss lessons learned in regard to principals' views of teacher evaluation and its implementation in schools. As such, both commonalities and differences will be examined in the next sections.

Putting policy into practice: commonalities and challenges from a principal perspective

In this section both commonalities and challenges in implementing teacher evaluation are identified taking a principal perspective. Strategies principals used to meet those challenges are also examined. In general, based on both studies' data, four key themes were identified: a) balancing conflicting goals of evaluation, b) minimising negative effects of evaluation on teachers, c) managing tensions of implementation, and d) making sense of the new policy and its effects at school. Table 2 summarises the main challenges and strategies identified.

INSERT TABLE 2 HERE

Balance conflicting goals

Principals identified a conflict between what they perceived as two separate purposes of evaluation: supporting formative professional growth and conducting summative evaluation judgments. Although the goals of the teacher evaluation systems are usually both developmental

and summative, at least at a rhetoric level, in practice the formative dimension tends to be subsumed by the summative purpose (Avalos, 2010; AUTHOR 2, 2009, 2010). Principals believed that evaluation's main goal should be to promote the professional growth and development of all teachers. Leading teacher professional growth was a responsibility they continued to embrace. In contrast, principals believed the mandated policy on teacher evaluation emphasised a bureaucratic approach and summative judgments, often perceived as unfair. These two differing and competing goals of teacher evaluation often resulted in conflicting choices for principals. For example, formative observation data collected over time was not always congruent with the final collection of evidence including student test scores that became part of the summative evaluation. For instance, a United States principal described the two beliefs in operation:

Now if I think you're [the teacher] doing something wrong, then I'm going to be the first to go to you and say, 'You need to try something different,' but it's always in an effort to have the teacher get better. I don't ever want to beat somebody over the head with anything like this [the evaluation system].

Portuguese principals also identified the incongruous purpose and reality regarding the system's intended outcome. As a Portuguese principal commented,

There is a mismatch between intentions at a rhetoric level and reality in terms of the purpose of the appraisal system. It is an administrative process. It does not lead to improving education or teachers' development.

Principals sought to avoid difficulty leading in systems with incompatible goals. Thus they merged and maintained a balance of professional development actions and summative judgment requirement. They continued to promote teacher growth as the main purpose of evaluation. At the same time they complied with policy directives and rated, ranked, and reported teacher competencies or lack of them. They focused on students and learning even when policies might not specifically address student learning. They actively sought to maintain a positive school climate by mediating negativity as the following comment from a Portuguese principal illustrates:

We have done what we could to avoid problems at school. We have negotiated and reflected with the teachers in order to avoid the negative effects that happened in other schools in order to maintain the good school atmosphere.

Principals took a strong role implementing the evaluation process in the schools. These actions held the school course steady and able to withstand the disruptive aspects of change.

Minimising negative effects

The teacher-principal relationship is central to the evaluation process, regardless of teacher evaluation policy mandates. In both studies, data indicated that i) principal- teacher relationships are important, ii) evaluation policies can affect relationships either negatively or positively, and iii) collaborative relationships are valued.

Valuing their relationships with teachers, in general, principals demonstrated strong collaborative actions with teachers as they learned the nuances of evaluation model together. One

US principal observed, ‘When we started out the year, it was new for all of us; so we worked off the theme of ‘If you hold my hand, I’ll hold yours.’

Further elaborating on the valued supportive relationship, another described how professional knowledge surpasses policy:

We know our teachers. We know what is normal or abnormal with their procedures and teaching styles. So if I walk into a classroom even if it is a planned observation, and this is just a really off day for them, I’m going to spend 10 or 15 minutes in there and I’m going to get up and leave and we’re going to do that over some other time. We all have off days. We’re not setting them up for failure. We don’t want a ‘gotcha’ game.

Paradoxically, less time with teachers in the United States study was an unintended consequence of the new policy. The new teacher evaluation system was intended to increase principal time in the classroom for observation and evaluation. However, the process required more time in the office writing evaluations and entering the observations scores into the computer. Moreover, principals indicated that important but less formal time with teachers suffered as well as their ability to monitor school safety, classroom management, and student welfare as indicated by the following comment:

The new evaluation process has actually taken us out of the classrooms because previously we were in there every day. That is a complaint from teachers too. We are not nearly as physically available [as before evaluation implementation].

Likewise, Portuguese principals valued the relationship between school leaders and teachers. Leadership practices maintained a steady course buffering perceived negative effects. Principals foreshadowed the potential of the new evaluation to negatively affect staff relationships. Some reports of the possible deterioration of professional collaborative relationships troubled principals. It was feared that rating and ranking scores or teachers might increase individualism and competitiveness. The emergence of tensions and conflicts, the lack of recognition, and a decrease in teacher collaboration are recurring themes in Portuguese principals' accounts and emerged as a worrisome possibility in the United States principal interviews. Issues of teacher participation and sense of ownership in regard to the new evaluation system as well as the mediating influence of school leadership in implementing it explained, at least in part, the emergence of conflicts and tensions in some contexts and the acceptance, or at least, the recognition of some of positive features of the new policy in other contexts. The following quotation illustrates the theme:

The school climate has changed due to greater tensions amongst teachers. Some teachers do not believe or don't care about the appraisal process and other teachers were really involved in it, so tensions and conflicts have emerged. And this was not the case before the implementation of the appraisal process.

Managing tensions of implementation

Principals reported implementation tensions and the themes emerged across the two studies. The tensions resulted from : i) a confusing and insufficiently developed policy model of

evaluation, and ii) increased workload resulting from adding the evaluation implementation to existing work, and iii) lack of training and time for implementation.

United States principals reported frustration with the new evaluation policy's complexity. For example, clear guidelines on reporting procedures from the state department were not available due to the hurried implementation. Furthermore, the newly legislated teacher evaluation process eliminated tenure, but specific details were not available. As a result, principals reported that monitoring the many unsubstantiated rumours took time away from the more positive aspects of the evaluation process as indicated in the following principal comment, 'Teachers were scared. It [evaluation implementation] could have been done in a less painful way.'

Portuguese principals described the dilemmas similarly as the following quote illustrates:

This appraisal system was a confusing and complex process. It was not transparent and it was poorly developed. It was too time-consuming especially as far as the bureaucratic work was concerned. And this has prevented teachers from their main role: teaching.

Overall, principals were sceptical about the evaluation system's effect on their school and teachers due mainly to a confusing or unclear model (see Table 3). As reflected in the following comments, they described the increased bureaucracy, teachers' lack of motivation, and the endless changes in legislation:

The main effects at my school are teacher lack of motivation and disillusionment. Teachers feel that it is a waste of time being involved in the appraisal system because they

believe it is mainly a bureaucratic and administrative device with no impact on improving their professional practice.

INSERT TABLE 3 NEAR HERE

Principals across the studies frequently indicated the tremendously time-consuming aspects of the implementation. Principal comments indicated concern with student outcomes: ‘Teacher energy is focused on the appraisal system rather than on student learning.’

Another theme emerging from the studies was that lack of training and support principals received hindered the implementation process. The United States principals reported that training was insufficient to learn about the significant changes in teacher evaluation and to implement them.

Lack of time for the implementation was the major concern for principals in both studies. The United States principals reported that implementing the new evaluation policies demanded a significant reorganisation and reprioritisation of their school day schedule. The focus on evaluation often resulted in the exclusion of other duties as illustrated by the following comment: ‘I’ve not done a real good job of balancing [numerous school duties]. In the first semester, for example, it was evaluations all the time. You put your head down and you went at it.’

Another concern was the impact on the school leaders’ personal time resulting from the increased time demands of implementing the teacher evaluation process as the following comment from a United States principal indicates: ‘It has taken a lot of time outside of working hours. I mean weekends and nights and breaks. You’ve got to find the hours somewhere and it’s hard to do. We don’t like it to take away from the school day.’ Similarly, Portuguese principals

agreed that time was not available for teacher evaluation responsibilities acknowledging that, ‘It is a very bureaucratic model that leads to an increase of the workload.’

Lastly, changing details of the policies in both studies resulted in even further demands on time as a Portuguese principal noted, ‘If the teacher appraisal is to lead to improvement, the endless changes in legislation do not allow any kind of improvement, as you have to start all over again all the time.’ Echoing this sentiment a United State principal commented stating ‘The state needs to stop moving the [policy] target.’ when discussing the frequent changes that required first learning and relearning the system requirements.

Making sense of the new policy and its effects at school

A step to understanding the meaning of the rubric’s elements many principals promoted collaborative teacher conversations during summative evaluation conferences. The principals reported that the rubric in the US context was instrumental in changing or enhancing their understanding of what good teaching looks like in the classroom. They reported that this knowledge assisted them in recognising numerous teaching competencies and then in providing teachers feedback on their performance. U.S. principals used the evaluation rubric as a vehicle to focus teacher conversation on learning as indicated by the following comment: ‘It’s a different way of teaching and so we’re having more conversations with our teachers.’

In addition to increased principal-to- teacher conversations, U.S. principals reported increased teacher-to-teacher collaborative dialogue as well. The need to understand and unpack the unfamiliar rubric’s meaning necessitated discussion. As one principal commented, ‘They [the

teachers] have really been able to work together as a team with that rubric.’ Another added, ‘The in-depth planning required for this rubric has forced them to even collaborate more because one person simply can’t do it all. There aren’t enough hours in the day.’ Yet another principal provided an example illustrating learning through teacher-initiated dialogue and discussion:

We sat down one afternoon and we went through it [the rubric]. We spent two hours after school. Then the teachers actually got together by themselves. I would go into rooms after school, and they would be there for an hour and a half, studying the rubrics.

On the other hand, Portuguese principals were more critical of the new policy and its effects on school development. In general, they spoke of either no effect or a negative effect on teachers’ professional development as a result of the new system’s implementation. When principals were asked if there was an effect on the school due to teachers’ performance appraisal, many responded that the effects were negative or no effect. Furthermore, they reported the evaluation had little or no effect on improving teaching quality. By and large, principals are sceptical about the appraisal system’s effect on their school and teachers. As reflected in the following comments, they described the increased bureaucracy, teachers’ lack of motivation, and the endless changes in legislation:

The main effects at my school are teacher lack of motivation and disillusionment.

Teachers also feel that it is a waste of time being involved in the appraisal system because they feel that it is mainly a bureaucratic and administrative device with no impact on improving their professional practice.

If the teacher appraisal is to lead to improvement, the endless changes in legislation do not allow any kind of improvement, as you have to start all over again all the time.

However, some principals claimed that it is up to them to make a difference in implementing the appraisal process. This attitude is reflected in the following response.

The school climate has been the same as before the implementation of the appraisal system and teachers keep on working in collaboration. We have done what we could to avoid problems at school. We have negotiated and reflected with the teachers in order to avoid the negative effects that happened in other schools in order to maintain the good school atmosphere.

Conclusion and discussion

The need to improve teaching quality and standards in order to increase student achievement has led governments worldwide to implement accountability policies, an example of which is teacher evaluation. Principal perception of a new teacher evaluation system is important to understand as the principal assumes the primary responsibility for its implementation at the school. Moreover, the principal's perception of the evaluation system's efficacy contributes to how policies are interpreted and put into place. Thus success or failure is dependent in part on the principal's beliefs, perceptions and strategies. Teacher collaboration as a key issue in teacher evaluation at school, as well as the development of rubrics that teachers rely on and the adequate training of appraisers are few examples of the ways in which school leadership has impacted upon the enactment of teacher evaluation policy in different schools. A great deal of similarity exists between principals' perceptions in the two studies on which this paper is drawn, indicating the possibility of a global, rather than just a local, concern. The

following sections discuss the similarities and any differences between the principal perceptions in both studies.

School leadership and supportive school structures for teacher evaluation

Implementing a new teacher evaluation policy is problematic (Timperley & Robinson, 1997). One significant problem is that principals' perceptions and experiences as appraisers mediate policy. Appraisers perceived effectiveness of an evaluation system has been linked to greater acceptance and commitment to the implementation (Taylor, 2010). A higher level of belief in the evaluation efficacy implies greater support for the implementation (Johnson, Reckers & Bartlett, 2014). A positive perception of the evaluation system is even more important for successful implementation when those responsible for implementation are not involved in its development (de Wall & Counet, 2009) as is the context of this study. Therefore, a better understanding of principals' experiences with and perceptions of evaluation is required (Ovando & Ramirez, 2007) if the success of a new system of teacher evaluation is the desired outcome.

Trust, vision, support and structure have been identified as key dimensions of school leadership influencing teachers' perceptions of an appraisal system (Tuytens & Devos, 2010). Also relevant is a constructive school climate in which to implement such a system (Stronge & Tucker, 2003). Ovando and Ramirez (2007) confirm that principals' significant instructional leadership actions include setting clear expectations, monitoring instruction through walk-through observations, and providing professional development opportunities based on teachers' needs. Those needs are critical, particularly when fostering the formative dimension of teacher

evaluation. In fact, research has emphasised the influence of school cultures and organisational features in teacher learning and professional development (AUTHOR 1, 2004; AUTHOR 1 & AUTHOR, 2014).

Drawing upon policy characteristics that Fullan (2001, 2007) identified as affecting policy implementation, Tuytens and Devos (2010) concluded in their study in Belgium that the structure a principal provides, along with the trust teachers have in their principal, is of central importance to teachers' perceptions of the policy's practicality. The same study also illustrated that vision is one of the most important dimensions influencing teachers' perceptions of the policy as it affects two essential policy characteristics: need and practicality. Empirical work has also demonstrated that leadership directly influences feedback's usefulness and, thus, indirectly influences teachers' professional learning (Tuytens & Devos, 2011).

School administrators acknowledge that performance appraisal is beneficial in goal setting. Likewise, enhanced supervision and communication have the potential for school improvement, including creating a common school-wide language for teaching and learning (Kersten & Israel, 2005). However, research literature also highlights constraints that might hinder the appraisal process, including resistance; lack of evaluation feedback; limited sharing of practice evaluators' lack of expertise; poor understanding of the appraisal's purposes; excessive workload; and lack of resources, including time (OECD, 2013).

Interpreting and managing policy in context: dealing with mediating factors

Implementing a new policy is a complex and dynamic process which is more problematic when the policy impacts school culture and structures and collaborative relationships. As key

players and mediators in policy implementation at schools, principals both influence and are influenced by teachers' perceptions about the new policy. Principals in these studies sought to minimise negative effects caused by the implementation of the externally mandated teacher evaluation. They strove to protect their relationship with teachers. Indeed, principal-teacher relationships were deemed to be as important and beneficial in the supervisory relationship. Thus, tasked with implementing an accountability system granting more direct authority over teachers, principals in these studies tried to keep positive relationships and teacher support at the centre of their philosophy. According to Lortie (2009), the relationship between principals and teachers both individually and collectively is "the most important priority for a principal to establish and sustain" (p.79). In general, the principals wanted to be seen as warm, caring, communicative leaders, worthy of teacher trust.

Principals expressed the need to balance conflicting goals as they maintained a collaborative and positive school culture yet at the same time comply with a potentially contentious mandate changing the process and content of teacher evaluation. As part of this supportive relationship, principals when implementing accountability-driven externally mandated evaluation policies perceived their responsibility to buffer and protect teachers from evaluation's negative effects. They attempted to learn the new system and apply it while tempering any negative effects on teachers with whom they feel a familial relationship (AUTHOR 2 & AUTHOR, 2013).

Enhancing the formative dimension of evaluation: the role of supervision

Principals clearly espoused a formative philosophy and viewed assisting teachers to improve as an important part of their job. Principals carefully balanced providing professional development to teachers with the evaluation system's summative mandate to rate and score them. They provided guidance and support to teachers even though the change caused strong emotional teacher reactions. When emotions are strong and differences of opinion are plentiful (Fullan, 2001), principals must work to maintain positive relationships with teachers while integrating the new evaluation policies, rubrics, and practises mandated by the state or national government.

Time is, and was, a resource in short supply. However, if implementation of a new teacher evaluation system is to succeed, time constraints must be mitigated. In both studies, lack of time was seen as a major hurdle. For principals, the combination of learning a dramatically different evaluation process and assisting teachers to learn the new system while also evaluating them was stressful. The time requirements can detract from fulfilling other important duties, including being present among students in the halls. When principals perceive a new evaluation system as a waste time spent on compliance rather than supporting teacher improvement, implementation with vigour or fidelity is compromised.

Interesting differences and implications for further studies

The use of an instructional rubric to promote instructional conversations between principal and teachers was a positive aspect in the US study. Principals embraced the detailed instructional rubric as a map or an explanation of what good teaching looks like. Prior to the introduction of the rubric, evaluation criteria were brief and not specific consisting of half a dozen general comments. In contrast, the new rubric had 92 elements of instruction that principals used for observations and subsequent feedback to teachers.

In Portugal the rubrics have been subject to changes as a result of changes in the legal framework along with the introduction of both external and internal teacher evaluation. The most critical issues related to the training of the appraisers (teachers who perform the role of appraisers) and the existence of a quota system.

A second difference between the two studies is in the perceived intensity of government involvement. While principal in both countries were sceptical of the external governmental mandates, Portugal produced significantly more legal documents over a greatly length of time. For 5 years, the Portugal evaluation system was subjected to changing legal requirements. In contrast, the US schools in this study had not seen a major change in the evaluation system for nearly a decade. Although the US change was significant, the positive aspect of the rubric was was important as a tool for principals to improve their skill in providing instructional supervision to teachers.

As teacher evaluation systems are redesigned in response to changing national and state policies, principals' perceptions in the schools we studied might help leaders design more effective implementation plans. Concerns raised in this study may also contribute to conversations among policy makers about potentially counterproductive effects on principals' and teachers' work as a result of underdeveloped and designed mandates. If higher student achievement and greater teacher effectiveness are to be realised, implementation problems and counter-productive system design features must be minimised.

This paper raises other questions that may be pursued through additional research: How have principals' perspectives changed over time? And why? Has teacher evaluation increased student achievement? Are more gradual cultural changes occurring in schools as a result of the new evaluation system? Furthermore, additional research may also probe the interaction of new

teacher evaluation systems with other concurrent teacher improvement initiatives, such as increased professional development for teachers and substantial curriculum changes.

References

AUTHOR 1 (2004).

AUTHOR 1 (2012).

AUTHOR 2 (2009)

AUTHOR 1 & AUTHOR (2014)

AUTHOR, 2 & AUTHOR. (2013)

AUTHOR 2 (2014)

Avalos, B. (2010). O sistema chileno de avaliação do desempenho docente. In: FLORES, M. A. (Org.). (2010). *A Avaliação de professores numa perspetiva internacional: sentidos e implicações*. Coleção Saberes Plurais. Porto: Areal Editores.

Barzano, G. & Grimaldi, E.(2013). Discourses of merit: The hot potato of teacher evaluation in Italy. *Journal of Education Policy*. 28:6,767-791.

Darling-Hammond, L. (2013). *Getting teacher evaluation right*. New York: Teachers College Press.

De Wall, A.,Counet, H. (2009). Lessons learned from performance management systems implementation. *International Journal of Productivity and Performance Management* ,58 (4), 367-390.

Fullan, M. (2001). *Leading in a culture of change*. San Francisco: Jossey- Bass.

Fullan, M. (2007). *The new meaning of educational change*. Amsterdam, NY: Teachers College Press.

- Johnson, E.N., Reckers, P.H., Bartlett, G.D. (2014). Influence of timeline and perceived strategy effectiveness on Balanced Scorecard performance evaluation judgments. *Journal of Management Accounting Research*, 26 (1), 165-184.
- Kersten, T. A., & Israel, M. S. (2005) Teacher evaluation: principals' insights and suggestions for improvement. *Planning and Changing*, 36(1&2), 47-67.
- Liu, S., & Zhao, D. (2013). Teacher evaluation in China; latest trends and future directions. *Educational Assessment, Evaluation and Accountability*, 25 (3), 321-250
- Lortie, D. (2009). *School principal*. Chicago: University of Chicago Press.
- Mango, C.S. (2013). *Comparative perspectives on international school leadership*. New York: Routledge.
- Marshall, K. (2009). *Rethinking teacher supervision and evaluation*. San Francisco: Jossey-Bass.
- OECD. (2009). *Evaluating and rewarding the quality of teachers: International practices*. Paris: OECD.
- [Organisation for Economic Co-operation and Development, OECD](#) (2013) *Teachers for the 21st century: Using evaluation to improve teaching*, OECD Publishing.
- Ovando, M.N., & Ramirez, A. (2007). Principals' instructional leadership within a teacher performance appraisal system: enhancing students' academic success. *Journal of Personnel Evaluation in Education*, 20 (1-2), 85-110.
- Retallick, J., & Fink, D. (2002). Framing leadership: contributions and impediments to educational change. *International Journal of Leadership in Education*, 5(2), 91-104.
- Stronge J.H., & Tucker, P.D. (2003). *Teacher evaluation. Assessing and improving performance*. Larchmont, NY: Eye on Education.
- Sun, M., Youngs, P., Yang, Chu, H. & Zhao, Q. (2012). Association of District Principal

Evaluation with learning-centered leadership practice: Evidence from Michigan and Beijing, *Educational Assessment, Evaluation and Accountability*, 24 (3), 189-213.

Taylor, W. (2010). The Balanced Scorecard as a strategy evaluation tool: The effects of implementation involvement and a causal-chain focus. *The Accounting Review*, 85 (3), 1095-1117.

Timperley, H.S., & Robinson, V.M.J. (1997). The problem of policy implementation: The case of performance appraisal. *School Leadership and Management*, 17(3): 333-345.

Tuytens, M., Devos, G. (2010). The influence of school leadership on teachers' perception of teacher evaluation policy. *Educational Studies*, 36(5), 521-536.

Tuytens, M., Devos, G. (2011) Stimulating professional learning through teacher evaluation: An impossible task for the school leader? *Teaching and Teacher Education*, 27(5), 891-899.

Vekeman, E.; Devos, G., & Tuytens, M. (in press) The influence of teachers' expectations on principals' implementation of a new teacher evaluation policy in Flemish secondary education, *Educational Assessment Evaluation and Accountability*, DOI: 10.1007/s11092-014-9203-4

Weisberg, D., Sexton, S., Mulhern, J. & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Retrieved from <http://widgeteffect.org/downloads/TheWidgetEffect.pdf>